

TTGTAACAGA AAATTAAT ATACTCCACT CAAGGGATT CTGTACTTGT CCCTTTGGT -99  
AAAGTCTCAT TTACATTCT AAACCTTCT TAAGAAAATC GAATTCCTT TGATCTCTCT -39

1 -1 M T S C H I  
TCTGAATTGC AGAAATCAGA TAAAAACTAC TTGGTGAA ATG ACT TCT TGT CAC ATT 18  
7 A E E H I Q K V A I F G G T H G  
GCT GAA GAA CAT ATA CAA AAG GTT GCT ATC TTT GGA GGA ACC CAT GGG 66  
23 N E L T G V F L V K H W L E N G  
AAT GAG CTA ACC GGA GTA TTT CTG GTT AAG CAT TGG CTA GAG AAT GGC 114  
39 A E I Q R T G L E V K P F I T N  
GCT GAG ATT CAG AGA ACA GGG CTG GAG GTA AAA CCA TTT ATT ACT AAC 162  
55 P R A V K K C T R Y I D C D L N  
CCC AGA GCA GTG AAG AAG TGT ACC AGA TAT ATT GAC TGT GAC CTG AAT 210  
71 R I F D L E N L G K K M S E D L  
CGC ATT TTT GAC CTT GAA AAT CTT GGC AAA AAA ATG TCA GAA GAT TTG 258  
87 P Y E V R R A Q E I N H L F G P  
CCA TAT GAA GTG AGA AGG GCT CAA GAA ATA AAT CAT TTA TTT GGT CCA 306  
103 K D S E D S Y D I I F D L H N T  
AAA GAC AGT GAA GAT TCC TAT GAC ATT ATT TTT GAC CTT CAC AAC ACC 354  
119 T S N M G C T L I L E D S R N N  
ACC TCT AAC ATG GGG TGC ACT CTT ATT CTT GAG GAT TCC AGG AAT AAC 402  
135 F L I Q M F H Y I K T S L A P L  
TTT TTA ATT CAG ATG TTT CAT TAC ATT AAG ACT TCT CTG GCT CCA CTA 450  
151 P C Y V Y L I E H P S L K Y A T  
CCC TGC TAC GTT TAT CTG ATT GAG CAT CCT TCC CTC AAA TAT GCG ACC 498  
167 T R S I A K Y P V G I E V G P Q  
ACT CGT TCC ATA GCC AAG TAT CCT GTG GGT ATA GAA GTT GGT CCT CAG 546  
183 P Q G V L R A D I L D Q M R K M  
CCT CAA GGG GTT CTG AGA GCT GAT ATC TTG GAT CAA ATG AGA AAA ATG 594  
199 I K H A L D F I H H F N E G K E  
ATT AAA CAT GCT CTT GAT TTT ATA CAT CAT TTC AAT GAA GGA AAA GAA 642  
215 F P P C A I E V Y K I I E K V D  
TTT CCT CCC TGC GCC ATT GAG GTC TAT AAA ATT ATA GAG AAA GTT GAT 690  
231 Y P R D E N G E I A A I I H P N  
TAC CCC CGG GAT GAA AAT GGA GAA ATT GCT GCT ATC ATC CAT CCT AAT 738  
247 L Q D Q D W K P L H P G D P M F  
CTG CAG GAT CAA GAC TGG AAA CCA CTG CAT CCT GGG GAT CCC ATG TTT 786  
263 L T L D G K T I P L G G D C T V  
TTA ACT CTT GAT GGG AAG ACG ATC CCA CTG GGC GGA GAC TGT ACC GTG 834  
279 Y P V F V N E A A Y Y E K K E A  
TAC CCC GTG TTT GTG AAT GAG GCC GCA TAT TAC GAA AAG AAA GAA GCT 882  
295 F A K T T K L T L N A K S I R C  
TTT GCA AAG ACA ACT AAA CTA ACG CTC AAT GCA AAA AGT ATT CGC TGC 930  
311 C L H .  
TGT TTA CAT TAG AA ATCACTTCCA GCTTACATCT TACACGGTGT CTTACAAATT 984  
CTGCTAGTCT GTAAGCTCCT TAAGAGTAGG GTTGTGCCTT ATTCAACTGC ATACATAGCT 1044  
CCTAGCACAG TGCCTTATTC GGTAGGCATC TAAGCAAATT TCTTAAATTAA ATTAATATAT 1104  
CTTTAAAGAT ATCATATTCTT ATGTATGTAG CTTATTCAA GAAAGTGTTC CTATTCTAT 1164  
ATAGTTTATT ATACATGATA CTTGGGTAGC TCAACATTCT TAATIAACAG CCTTTGTATT 1234  
CAGAATATAA AATTGAAATA GATATATATA AAGTTAAAAA AAAAAAAA AAA 1287

Fig. 1

	10v	20v	30v	40v	50v
HLASP	MTSCHIAEEHIQKVAIFGGTHGNELTGVFLVKHWLENGAEIQRTGLEVKPF				
	MTSCH:AE:I:KVAIFGGTHGNELTGVFLVKHLEN::EIQRTGLEVKPF				
BASPCDNA	MTSCHVAEDPIKKVAIFGGTHGNELTGVFLVKHWLENSTEIQRTGLEVKPF				
	10^	20^	30^	40^	50^
	60v	70v	80v	90v	100v
HLASP	ITNPRAVKKCTRYIDCDLNRFIDLENLGKKMSEDLPYEVRAQEINHLFGP				
	ITNPRAVKKCTRYIDCDLNR:FD ENLGKK. SEDLPYEVRAQEINHLFGP				
BASPCDNA	ITNPRAVKKCTRYIDCDLNRFDPENLGKKMSEDLPYEVRAQEINHLFGP				
	60^	70^	80^	90^	100^
	110v	120v	130v	140v	150v
HLASP	KDSESYDIIFDLHN'TTSNMGCTLILEDSDRNNFLIQLMFHYIKTSAPLPCY				
	KDSESYDIIFDLHN'TTSNMGCTLILEDSDRN:FLIQLMFHYIKTSAPLPCY				
BASPCDNA	KDSESYDIIFDLHN'TTSNMGCTLILEDSDRNDFLIQLMFHYIKTSAPLPCY				
	110^	120^	130^	140^	150^
	160v	170v	180v	190v	200v
HLASP	VYLIEHPSLKYATTRSIAKYPVGIEVGPQPQGVLRADILDQMRKMIHALD				
	VYLIEHPSLKYATTRSIAKYPVGIEVGPQPQGVLRADILDQMRKMI:HALD				
BASPCDNA	VYLIEHPSLKYATTRSIAKYPVGIEVGPQPQGVLRADILDQMRKMIQHALD				
	160^	170^	180^	190^	200^
	210v	220v	230v	240v	250v
HLASP	FIHHFNEGKEFPPCAIEVYKIIEKVDYPRDENGEIAAAIIHPNLQDQDWKPL				
	FIH:FNEGKEFPPCAIEVYKI:KVDYPR:E:GEI:AAIHP:LQDQDWKPL				
BASPCDNA	FIHNFNEGKEFPPCAIEVYKIMRKVDYPRNESGEISAAIIHPKLQDQDWKPL				
	210^	220^	230^	240^	250^
	260v	270v	280v	290v	300v
HLASP	HPGDPMFLTLDGKTIPLGGDCTVYPVFVNNEAAYYEKKEFAKTTKLTLNAK				
	HP.DP:FLTDGKTIPLGGD TVYPVFVNNEAAYYEKKEFAKTTKLTLNA:				
BASPCDNA	HPEDPVFLTLDGKTIPLGGDVTYPVFVNNEAAYYEKKEFAKTTKLTLNAN				
	260^	270^	280^	290^	300^
	310v				
HLASP	SIRCCLH				
	SIR..LH				
BASPCDNA	SIRSSLH				
	310^				

Fig. 2

09065807 100104

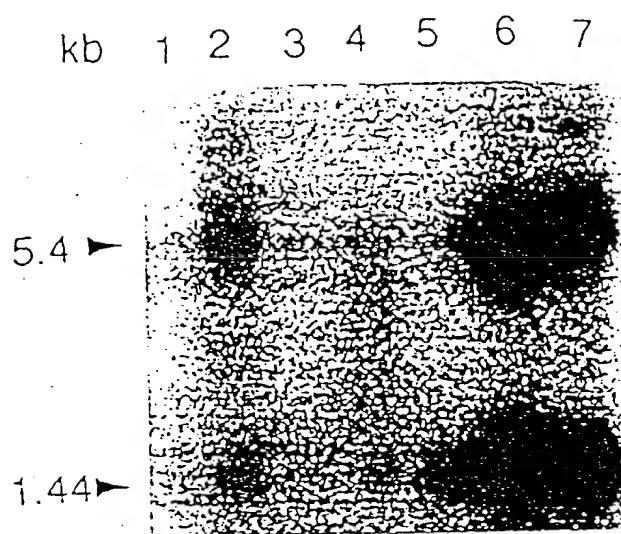


Fig. 3

Kaul et al.

165288-Z-000000

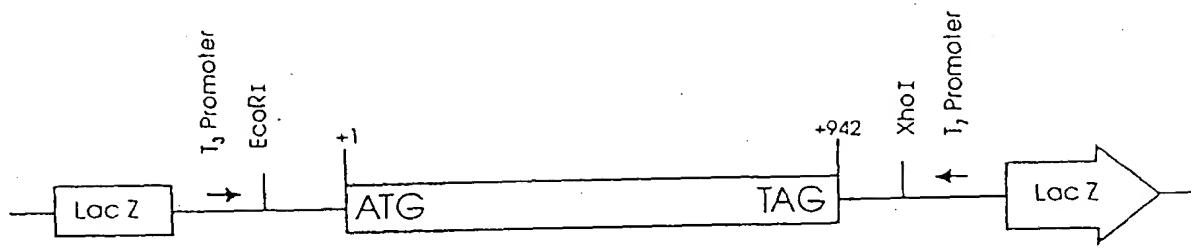


Fig. 4  
Kaul et al.

TOEPLITZ & SINGH

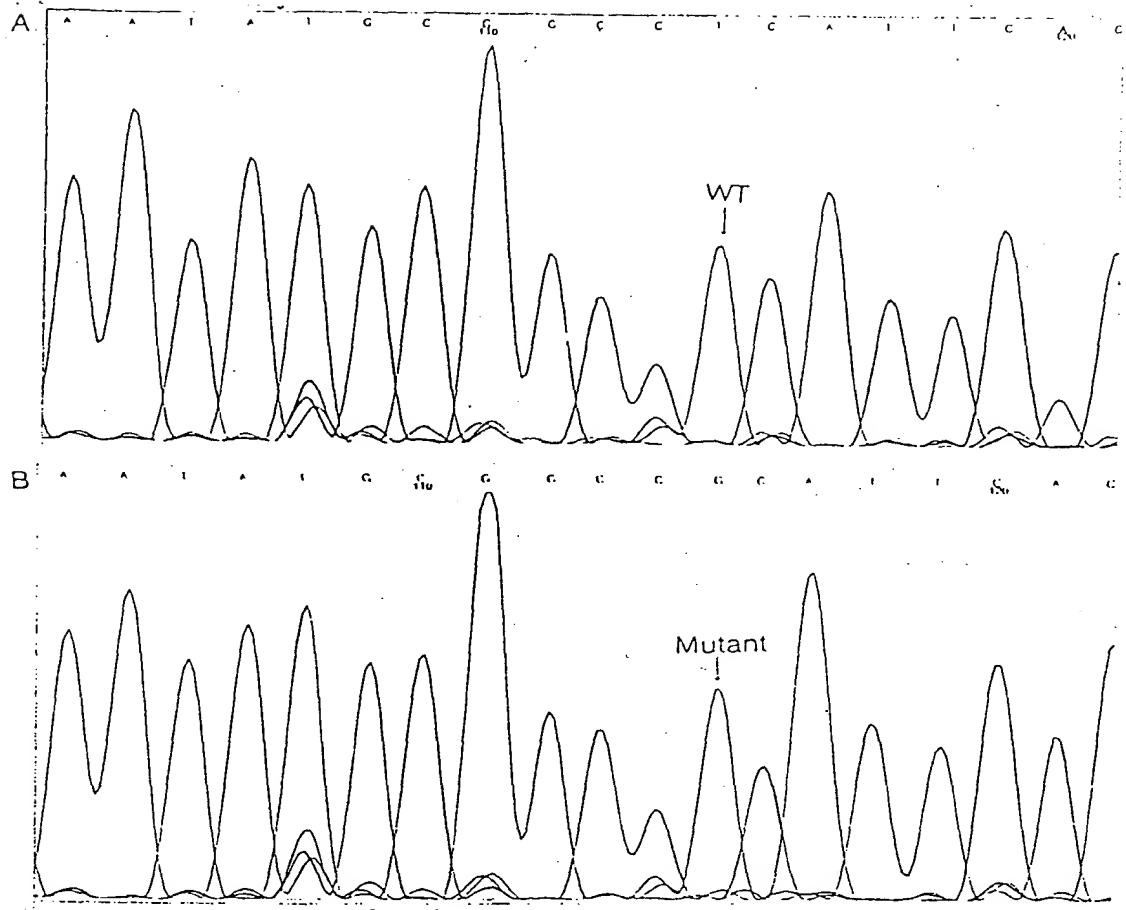
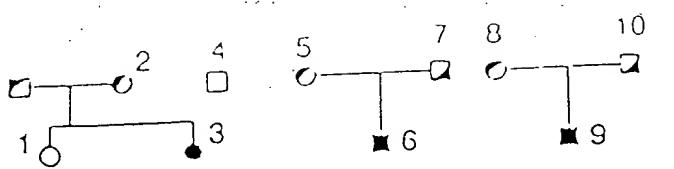


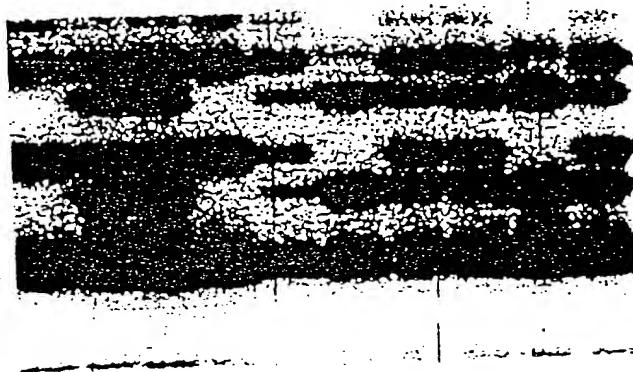
Fig. 5

Kaul et al.



1 2 3 4 5 6 7 8 9 10

A



B

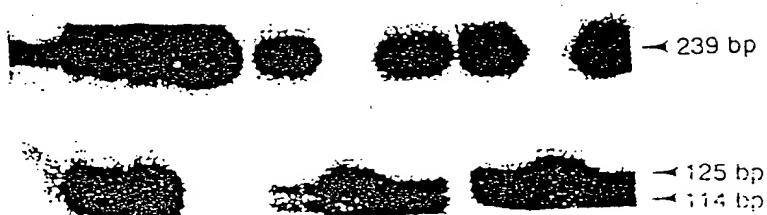


Fig. 6

Kaul et al.

8/26/93

MAPSEQ V5.33 HASP.SEQ(1,1277) Reading frames: 1 Enzyme file ALL.ENZ LinPage 1

EAM	M	E	M	N	DSNDSBB
ASA	B	C	N	L	STCSESS
MPE	O	5	L	A	AYOACAA
1E3	2	7	1	4	11111JJ

/

ATGACTTCTTGTACATTGCTGAAGAACATATACAAAAGGTTGCTATCTTGAGGAACC

-----+-----+-----+-----+-----+-----+-----+-----+ 60  
TACTGAAGAACAGTGTAAACGACTTCTTGTATATGTTTCCAACGATAGAACCTCCTTGG

m t s c h i a e e h i q k v a i f g g t

-----+-----+-----+-----+-----+-----+-----+-----+

N	A	BBH	TSM	RM	H	HHD	TH
L	L	SCP	RPS	MA	I	HAD	FN
A	U	AAA	UOE	AE	N	AEE	IF
3	1	W72	911	11	P	121	11

/ / / / / / / /

CATGGGAATGAGCTAACCGGAGTATTCTGGTTAAGCATTGGCTAGAGAATGGCGCTGAG

-----+-----+-----+-----+-----+-----+-----+-----+ 120  
GTACCCTTACTCGATTGGCCTCATAAAGACCAATTCTGTAACCGATCTTACCGCGACTC

h g n e l t g v f l v k h w l e n g a e

-----+-----+-----+-----+-----+-----+-----+-----+

M	B
N	P
L	M
1	1

ATTCAGAGAACAGGGCTGGAGGTAAAACCATTATTACTAACCCCAGAGCAGTGAAGAAG

-----+-----+-----+-----+-----+-----+-----+-----+ 180  
TAAGTCTCTTGTCCGACCTCCATTGGTAAATAATGATTGGGTCTCGTCACTTCTTC

i q r t g l e v k p f i t n p r a v k k

-----+-----+-----+-----+-----+-----+-----+-----+

*Restriction Sites*

1 - 67 15  
 68 - 131 16  
 132 - 141 17

Fig. 7(a)

SEQ V5.33 HASP.SEQ(1,1277) Reading frames: 1 Enzyme file ALL.ENZ LinPage 2

CR	M	M	TH
SS	B	A	FN
PA	O	E	IF
61	2	3	11

TGTACCAGATATATTGACTGTGACCTGAATCGCATTGGACCTTGAACAAATCTTGGCAAA  
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 240  
ACATGGTCTATATAACTGACACTGGACTTAGCGTAAAAACTGGAACCTTTAGAACCGTTT

c t r y i d c d l n r i f d l e n l g k  
-----+-----+-----+-----+-----+-----+-----+-----+

NM	BN
DB	AS
EO	NP
12	22

/ AAAATGTCAGAAGATTGCCATATGAAGTGAGAAGGGCTCAAGAAATAATCATTATTT  
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 300  
TTTTACAGTCTTCTAACGGTATACTCACTCTTCCCAGTTCTTATTTAGTAATAAAA

k m s e d l p y e v r r a q e i n h l f  
-----+-----+-----+-----+-----+-----+-----+-----+

A	TH	M	S
V	FN	B	P
A	IF	O	O
2	11	2	1

/ GGTCCAAAAGACAGTGAAGATT CCTATGACATTATTTGACCTTCACAAACACCACCTCT  
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 360  
CCAGGTTTCTGTCACTTCTAAGGATACTGTAATAAAACTGGAAGTGTGTGGTGGAGA

g p k d s e d s y d i i f d l h n t t s  
-----+-----+-----+-----+-----+-----+-----+-----+

Fig. 7(b)

MN A	HBN M	TH E AS	TM
NL P	GSS N	FN C PC	RS
LA L	IIP L	IF R YR	UE
13 1	AH2 1	11 2 11	91
/	//	/ /	/

AACATGGGGTGCACTCTTATTCTTGAGGATTCCAGGAATAACTTTAATTCAGATGTTT  
-----+-----+-----+-----+-----+-----+-----+-----+-----+  
TTGTACCCCACGTGAGAATAAGAACTCCTAAGGTCTTATTGAAAAATTAAGTCTACAAA

420

n m g c t l i l e d s r n n f l i q m f  
-----+-----+-----+-----+-----+-----+-----+-----+

TM	N	M F	E
RS	L	A O	C
UE	A	E K	O
91	4	2 1	B

CATTACATTAAAGACTTCTCTGGCTCCACTACCCTGCTACGTTATCTGATTGAGCATCCT  
-----+-----+-----+-----+-----+-----+-----+-----+  
GTAATGTAATTCTGAAGAGACCGAGGGTATGGGACGATGCAAATAGACTAACCTCGTAGGA

480

h y i k t s l a p l p c y v y l i e h p  
-----+-----+-----+-----+-----+-----+-----+-----+

S	M	A
F	N	V
A	L	A
N	1	2

TCCCTCAAATATGCGACCACCTCGTCCATAGCCAAGTATCCTGTGGTATAGAAGTTGGT  
-----+-----+-----+-----+-----+-----+-----+-----+  
AGGGAGTTATACGCTGGT GAGCAAGGTATCGGTTCATAGGACACCCATATCTTCAACCA

540

s l k y | a t t r s i a k y p v g i e v g  
-----+-----+-----+-----+-----+-----+-----+-----+

Fig.7(c)

SEQ V5.33 HASP.SEQ(1,1277) Reading frames: 1 Enzyme file ALL.ENZ LinPage 4

D	M	M D	A	E	BMDD	TM	115 4p
D	N	N D	L	C	IBPP	RS	
E	L	L E	U	R	NONN	UE	
1	1	1 1	1	V	1121	91	

CCTCAGCCTCAAGGGGTTCTGAGAGCTGATATCTGGATCAAATGAGAAAAATGATTAAA

600

GGAGTCGGAGTTCCCCAAGACTCTCGACTATAGAACCTAGTTACTCTTTACTAATTT

p q p q g v l r a d i l d q m r k m i k

-----+-----+-----+-----+-----+-----+-----+-----+

NN

HMHM

SL

INHN

PA

NLAL

H3

P111

/

CATGCTCTGATTTATACATCATTCAATGAAGGAAAAGAATTCCTCCCTGCCATT

660

GTACGAGAACTAAAATATGTAGTAAAGTTACTCCTTTCTTAAGGAGGGACGCCGTAA

h a l d f i h h f n e g k e f p p c a i

-----+-----+-----+-----+-----+-----+-----+-----+

E

BSBNXSASSBBHNSB FF F IF

C

SESCMMVCESSPCCB OO O TN

P

ACAIAAARCAAIRV KK K AU

1

J1J11111JJ2111 11 1 1H

545p

/. / / / / / / / / / /

GAGGTCTATAAAATTATAGAGAAAGTTGATTACCCCCGGGATGAAAATGGAGAAATTGCT

720

CTCCAGATATTTAATATCTCTTCAACTAATGGGGGCCCTACTTACCTCTTAACGA

e v y k i i e k v d y p r d e n g e i a

-----+-----+-----+-----+-----+-----+-----+-----+

c 693>a  
Y231>x

Fig. 7 (d)

S	PBMDD	F	ESASBBSBXBNMDDB
F	SIBPP	O	CEPCSSFIHALBPPI
C	TNONN	K	RCYRAAAANOMAONNN
1	11121	1	2111JJN12141211

////

GCTATCATCCATCTTAATCTGCAGGATCAAGACTGGAAACCACTGCATCCTGGGGATCCC  
-----+-----+-----+-----+-----+-----+-----+-----+ 780  
CGATAGTAGGTAGGATTAGACGTCCTAGTTCTGACCTTGCTGACGTAGGACCCCTAGGG

a i i h p n l q d q d w k p l h p g d p

-----+-----+-----+-----+-----+-----+-----+-----+

(9° 4f ✓

N	TM	B	MDBBBDBMA	BBAB	CR	CR
L	RS	B	BPBSPIBL	SSCS	SS	SS
A	UE	V	ONSCUNNOW	ILIM	PA	PA
3	91	2	121911122	Y112	61	61

/ / / / / / /

ATGTTTTAACTCTGATGGGAAGACGATCCACTGGGCGGAGACTGTACCGTGTACCCC  
-----+-----+-----+-----+-----+-----+-----+  
TACAAAAATTGAGAACTACCCTCTGCTAGGGTGACCDGCCTCTGACATGGCACATGGGG

m f l t l d g k t i p l g g d c t v y p

-----+-----+-----+-----+-----+-----+-----+-----+

SM	HIFA	H A
PN	ATNC	I L
OL	EAUI	N U
11	31H1	3 1

/ \* //  
GTGTTTGTGAATGAGGCCGCATATTACGAAAAGAAAGAAGCTTTGCAAAGACAAC TAAA  
-----+-----+-----+-----+-----+-----+-----+-----+ 900  
CACAAACACTTACTCCGGCGTATAATGCTTTCTTCGAAAACGTTCTGTTGATTT

v f v n e a a y y e k k e a f a k t t k

-----+-----+-----+-----+-----+-----+-----+-----+

AS:54>C  
E285>A

Fig. 7(e)

S	B	HIF	E	A
P	B.	NTN	C	L.
O	V	FAU	1	U
1	1	31H	5	1

Alfa 365>c  
C9L4>A

CTAACGCTCAATGCAAAAGTATTGCTGCTGTTACATTAGAAATCACTCCAGCTTAC  
-----+-----+-----+-----+-----+-----+-----+  
GATTGCGAGTTACGTTTCATAAGCGACGACAAATGTAATCTTAGTGAAGGTCGAATG

960

l t l n a k s i r c c l h . k s l p a y

-----+-----+-----+-----+-----+-----+-----+

RM	A	ATM
MA	L	FRS
AE	U	LUE
11	1	291

ATCTTACACGGTGTCTTACAAATTCTGCTAGTCTGTAAGCTCCTTAAGAGTAGGGTTGTG  
-----+-----+-----+-----+-----+-----+-----+  
TAGAATGTGCCACAGAACATGTTAAGACGATCAGACATTGAGGAATTCTCATCCAACAC

1020

i l h g v l q i l l v c k l l k s r v v

-----+-----+-----+-----+-----+-----+-----+

B	A	RM	H	D	S
S	L	MA	N	D	F
P	U	AE	F	E	A
W	1	11	3	1	N

CCTTATTCAACTGCATACATAGCTCCTAGCACAGTGCCTTATTGGTAGGCATCTAAGCA  
-----+-----+-----+-----+-----+-----+-----+  
GGAATAAGTTGACGTATGTATCGAGGATCGTCACGGAATAAGCCATCCGTAGATTG

1080

p y s t a y i a p s t v p y s v g i . a

-----+-----+-----+-----+-----+-----+-----+

Fig. 7(f)

SEQ V5.33 HASP.SEQ(1,1277) Reading frames: 1 Enzyme file ALL.ENZ LinPage 7

TM	ATM	PATM	TDM	E	A
RS	SRS	ASRS	RRS	C	L
UE	EUE	CEUE	UAE	R	U
91	191	1191	911	V	1
/	//	///	/		

AATTTCTTAAATTAAATTATATCTTAAAGATATCATATTTATGTATGTAGCTTATT  
-----+-----+-----+-----+-----+-----+-----+-----+  
TTAAAGAATTAAATTATAGAAATTCTATAGTATAAAACATACATCGAATAA

1140

n f l n . l i y l . r y h i l c m . l i

-----+-----+-----+-----+-----+-----+-----+-----+

X	N	A
M	L	L
N	A	U
1	3	1

CAAAGAAGTGTTCCTATTCCTATAGTTATTACATGATACTGGGTAGCTCAACA  
-----+-----+-----+-----+-----+-----+-----+  
GTTTCTTCACAAAGGATAAAGATATATCAAATAATATGTACTATGAACCCATCGAGTTGT

1200

q r s v s y f y i v y y t . y l g s s t

-----+-----+-----+-----+-----+-----+-----+-----+

✓	✓
TM	TM
RS	RS
UE	UE
91	91
/	/

TTCTTAATAAACAGCCTTGATTAGAATATAAAATTGAAATAGATATATAAAAGTTA  
-----+-----+-----+-----+-----+-----+-----+  
AAGAATTATTGTCGGAAACATAAGTCTTATTTAACCTTATCTATATATTCAAT

1260

f l i n s l c i q n i k l k . i y i k l

-----+-----+-----+-----+-----+-----+-----+-----+

AAAAAAAAAAAAAAA

-----+-----+-----+-----+-----+-----+-----+-----+  
1277

TTTTTTTTTTTTTTTT

k k k k k k

-----+-----+-----+-----+-----+-----+-----+-----+

Fig. 7(g)